Inference in Stochastic Frontier Models Based on Asymmetry

Ahmed S[1], Sonia Pérez-F[1], Carlos Carleos A[1], Norberto C[1] and Pablo Martínez C[2]*

[1]*Department of Statistics and Operational Research and Mathematics Didactics, University of Oviedo, Spain*

[2]*The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, USA*

Submission: September 18, 2016; **Published:** January 26, 2018

*Corresponding author: Pablo Martínez Camblor, The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Hanover, NH, USA, Email: Pablo.Martinez.Camblor@Dartmouth.edu

Abstract

Stochastic frontier analysis (SFA) is often employed to study the production functions. The structure of errors is the main difference between the standard regression analysis and the stochastic frontier models; in the SFA, an independent random term with positive value is added to the usual white noise error. Conventionally, the parameters involved in the SFA are estimated and then, the convenience of using this model is tested. The authors propose to study, previously, the residuals in order to check the capacity of assuming a stochastic frontier model and then, if applicable, to estimate the parameters. With this goal, several non-parametric hypothesis testing are explored. Monte Carlo simulations suggest that, under the usual conventions, simple inference based on skewness provided competitive results and proved to be robust in the presence of outliers. Standard methods based on the maximum-likelihood obtained really poor results; they only were reasonable when both the sample size and the inefficiency term were large.

Keywords: Confidence interval; Contaminated distribution; Hypothesis testing; Stochastic frontier analysis

Abbreviations: SFA: Stochastic Frontier Analysis; MOM: Method of Moments; ML: Maximum-likelihood; WHO: World Health Organization

Introduction

The main objective of the production frontier or Stochastic Frontier Analysis (SFA) is the study of the production function. In spite of the study of the empirical estimation of production functions have begun long before, the contemporary formulation of the problem was proposed by Aigner, et al. [1] and Meeusen, et al. [1]. Formally, the following general model is considered

$$Y = g(\mathbb{X}) + \varepsilon$$

Where, $\mathbb{X} = \{X_1, \cdots, X_k\}$ is the vector of regressors, $g(\cdot)$ is the deterministic production function and ε is the deviation from the frontier production. In the SFA model ε is considered to be $\varepsilon = \grave{o} - u,$ where $\grave{o}$ is the the usual white noise and u is a non-negative quantity which represents the {technical inefficiency}. Conventionally, both quantities are assumed to be independently drawn. Although different semi-parametric [3,4] and non-parametric models have been proposed for the production function [5], the translog formulation is still the most frequent,

$$Y = \beta_0 + \beta_1 \cdot X_1 + \cdots + \beta_k \cdot X_k + \grave{o} - u, \quad (1)$$

It is linear in parameters and the variables $(X_i, 1 \le i \le k)$ are the logarithm transformations of the originals. This specification has been shown mis-specify cost and production functions in many cases [6]; of course, a mis-specified response function is almost sure to affect estimated residuals. Since, the original normal-half normal and normal-exponential models were introduced by Aigner, et al. [1], different parametric models have been considered for the pair $\{\grave{o}, u\}$: Stevenson [7] introduced a normal-truncated normal and Greene [8] the normal-gamma models, other possibilities like Laplace-exponential, Cauchy-half Cauchy or normal-uniform have also been studied, reports related with these models can be easily found from direct internet research. The selected model can have relevance in the obtained results [9]; the original proposed models: normal-exponential and normal-half normal, are still the most used in practice. Of course, non-parametric methods have also been considered; see, for instance, Simar et al. [10] and references therein.

Maximum-likelihood (ML) based methods are typically used in order to estimate the parameters involved in both the production function and the error components [11]. However, in Monte Carlo simulation studies, it has been observed that the method of moments (MOM) can lead to minor mean square errors that ML for several particular case (2008) [12,13]. The

usual procedure estimates - when it is possible- the specific frontier parameters and then computes the *p-values.* When the parameters cannot be computed, it is assumed that the stochastic frontier model (SFM) is not adequate. In spite of the huge number of published works dealing with frontier models and related problems, hypothesis testing has been less considered in the specialized literature: via Monte Carlo simulations, Coelli [11] concluded that both the likelihood-ratio and Wald based tests obtained incorrect size while method based on the third moment of the ordinary square error regression obtained an adequate size; recently, Simar & Wilson [14] proposed to use bootstrap method in order to develop inferences for the parameters involved in the SFM; their method provides useful information about inefficiency and model parameters irrespective of whether residuals have the skewness in the desired direction; Kuosmanen & Fosgerau [15] proposed to evaluate the SFM validity from the skewness of the regression residuals.

Once the regression parameters of equation (1) have been estimated, the main difference between the stochastic frontier and the usual linear regression models is the residuals structure. Notice that, the given an adequate dataset, the observed residuals

are not drawn from ε but from $\varepsilon - \mathbb{E}[\varepsilon](= \varepsilon + \mathbb{E}[u])$. Hence, the (expected) observed effect of the so-called technical inefficient term is a slight asymmetry on the residuals. In this work, and with a similar approach to Kuosmanen & Fosgerau [15], the authors propose to use residuals skewness in order to test where this asymmetry can be detected. With this goal two difference measures are considered: on other hand, one based

on $\mathcal{P}\{\varepsilon \leq -\mathbb{E}[u]\}$ and, on the other hand, one based on the

third central moment, $\mathcal{M}_3 = \mathbb{E}\left[\left(\varepsilon - \mathbb{E}[\varepsilon]\right)^3\right]$. Additionally, the

relationship between $\mathcal{P}\{\varepsilon \leq -\mathbb{E}[u]\}$ and the inefficiency term is used in order to propose a slight modification on the MOM procedure which, via Monte Carlo simulations, is proved to be more robust than the rest of the studied methods in the presence of outliers.

Rest of the paper is organized as follows: section 2 is devoted to point out some brief remarks about the available software. Theoretical aspects of the proposed measures are exposed in section 3. Performance of four different tests are studied via Monte Carlo simulations in section 4; three different scenarios were considered: normal-half normal, contaminated normal-half normal and the well-known WHO data structure with normal-half normal residuals. In section 5 the WHO data are directly analyzed and, finally, in section 6, the obtained results are discussed.

Software

As in the rest of scientific fields, statistical advances are closely related with the computational developing. In addition,

the popularization of a particular statistical procedure strongly depends on the characteristic of the available software. Currently, there exist several programs and routines for stochastic frontier and efficiency analysis, the most general being LIMDEP/NLOGIT and Stata. In addition, the freeware program, FRONTIER 4.1, developed by Tim Coelli (it can be easily found on the web) can also be used for a small range of stochastic frontier models. Unfortunately, FRONTIER is quite old; however, a version of FRONTIER 4.1 in the free environment of R has been developed by Arne Henningsen. In addition, Belotti et al. have written some new routines for Stata. Christopher Parmeter has written some additional code in R. LIMDEP has included an extensive package for frontier modeling since the mid 1980s. New features and models are added to LIMDEP on an ongoing basis. Recently, Wilson [16] developed the R package FEAR for computing non-parametric efficiency estimates, making inference, and testing hypotheses in frontier models.

Inference based on asymmetry

Hypothesis testing: There are different ways to parameterize the stochastical frontiers models, among the most commonly used

terms are $\lambda = \sqrt{\mathbb{V}[u]/\mathbb{V}[\grave{o}]}$ and $\gamma = \mathbb{V}[u]/\left(\mathbb{V}[u] + \mathbb{V}[\grave{o}]\right)$ (in the normal-half normal model; where u follows a $\mathcal{N}^+(0, \sigma_u), \mathbb{V}[u]$, is replaced by σ_u^2. Notice that, due to the fact that λ relates the variance of the perturbance from the inefficiency term (u) with the variance of the white noise perturbance $(\grave{o})$, not with the total production measured; its value is not directly associated with the production inefficiency but with the global system inefficiency; i.e., the value of λ is affected by the capacity of the production function to predict the optimal production. Hence, in order to determinate the pertinence of applying an SFA, previously we must check if the available data provided enough statistical power to perform the testing,

$$\begin{cases} H_0 : \lambda = 0 \\ H_1 : \lambda > 0 \end{cases} \quad (2)$$

In practice, model (1) is transformed to obtain a new model in which, as usual, a zero-mean perturbation is involved,

$$Y = \beta_0 + \beta_1 \cdot X_1 + \cdots + \beta_k \cdot X_k + \grave{o} - u$$
$$= \alpha_0 + \beta_1 \cdot X_1 + \cdots + \beta_k \cdot X_k + \grave{o} - \left(u - \mathbb{E}[u]\right) \quad (3)$$

Where, $\alpha_0 = \beta_0 - \mathbb{E}[u]$. Both least square error (LSE) and ML methods perform adequately on this model. In fact, both procedures provide similar and good estimations for the parameters α_0, $\beta_i (1 \leq i \leq k)$ and $\mathbb{V}[\varepsilon] = \mathbb{V}[\grave{o}] + \mathbb{V}[u]$. Hence, in practice, since the above parameters have been

estimated, the obtained residual, $(\varepsilon =)\{e_1,\cdots,e_N\}$ (Nstands for the available sample size), can be considered as a random sample drawn from $\bar{\varepsilon} = \varepsilon + \mathbb{E}[u]$. Assuming that, under the null $(\lambda = 0)$, $\bar{\varepsilon}$ is symmetrically distributed, a positive $\lambda - value$ should produce a decreased on the probability that $\bar{\varepsilon}$ takes negative values. Note that this probability can be directly estimated from $\widehat{\mathcal{P}}_N^0 = N^{-1}\sum_{i=1}^N I_{(-\infty,0]}(e_i)$ $I_A(x) = 1$ if $x \in A$ and 0 otherwise). Therefore, applying the traditional sign test, we first propose to check the contrast in (2) from

$$\begin{cases} H_0 : \mathcal{P}\{\bar{\varepsilon} \le 0\} = 1/2 \\ H_1 : \mathcal{P}\{\bar{\varepsilon} \le 0\} < 1/2 \end{cases} \qquad (4)$$

Despite the dependence among elements of ε because $\varepsilon = Q_X y$ where y is the vector of N observations from Y and Q_X is the orthogonal subspace projection matrix over the N observations of $\mathbb{X}$, under the null, $N \cdot \widehat{\mathcal{P}}_N^0$ could be approximated by a $\mathcal{B}(N, 1/2)$ (binomial distribution with parameters N and $1/2$ and then, for a fixed nominal level α,

$$\mathcal{R}_P(\alpha) = \left\{\varepsilon \in \mathbb{R}^N \Big/ \widehat{\mathcal{P}}_N^0 < b_{N,\alpha}\right\} \qquad (5)$$

Where, $\mathcal{P}\{\mathcal{B}(N,1/2) < N \cdot b_{N,\alpha}\} \le \alpha$, is an approximated critical region for the contrast (4) at level α.

When, as usual, it is assumed that $\dot{o}$ follows a symmetric distribution $(\mathbb{E}[\dot{o}^3] = 0)$, it is also derived that $\mathcal{M}_3 = \mathbb{E}[\varepsilon^3] = \mathbb{E}[-u^3]$. Since, in the stochastic frontier model context $\lambda > 0, (\mathcal{M}_3 < 0)$, the contrast in (2) can also be tested from the contrast

$$\begin{cases} H_0 : \mathcal{M}_3 = 0 \\ H_1 : \mathcal{M}_3 < 0 \end{cases} \qquad (6)$$

Besides, for residuals normally distributed, the Central Limit Theorem (CLT) and the Slutski's Lemma guarantee the following result.

Theorem 1: Let $\{e_1,\cdots,e_N\}$ be a random sample from $\bar{\varepsilon}(\sim \mathcal{N}(0,\sigma))$. Then, if $\widehat{\mathcal{M}}_3 = N^{-1}\sum_{i=1}^N e_i^3$ and $\mathcal{M}_3 = \mathbb{E}[\bar{\varepsilon}^3]$,

$$\sqrt{N} \cdot \frac{\widehat{\mathcal{M}}_3 - \mathcal{M}_3}{\sqrt{6 \cdot \widehat{\sigma}^6}} \xrightarrow{\ \mathcal{L}\ } \mathcal{N}(0,1) \qquad (7)$$

Where, $\widehat{\sigma}(= \widehat{\sigma}(N))$ is a estimator for σ satisfying $(\widehat{\sigma} - \sigma) \xrightarrow{N} 0$ in probability. Hence, the set

$$\mathcal{R}_A(\alpha) = \left\{\varepsilon \in \mathbb{R}^N \Big/ \widehat{\mathcal{M}}_3 < -\lambda_\alpha \cdot \sqrt{6 \cdot \widehat{\sigma}^6 / N}\right\} \qquad (8)$$

Where, $\mathcal{P}\{\mathcal{N}(0,1) < -\lambda_\alpha\} \le \alpha$ is an (asymptotic) critical region for the contrast (6) at level α. It is worth to note that the asymptotic normality for the estimator $\widehat{\mathcal{M}}_3$ can be proved under weaker conditions; however, parametric models must be assumed in order to approximate its variance.

Confidence intervals

Once it has been decided that, for the studied dataset, a stochastic frontier is the appealing appropriate model (this decision can be taken even when the null (2) is not rejected; in fact, in order to estimate the involved parameter, we only need that $\widehat{\mathcal{P}}_N^0 < 1/2$ and $\widehat{\mathcal{M}}_3 < 0$; the question of the "wrong" skewness deserves some discussion which was included in final section), the next step is to estimate the involved parameters and computing the respective confidence intervals. Usually, some parametric model is assumed at this point. In this section, we use the $\mathcal{P}\{\bar{\varepsilon} \le 0\}$ properties to make inferences on λ for the normal-half normal model and, from Monte Carlo simulations, it is studied the behavior of the obtained expressions.

Under the normal-half normal model, it is well-known that, for each $x \in \mathbb{R}$, the distribution function of the composite disturbance $\dot{o} - u$ is,

$$\mathcal{P}\{\dot{o} - u \le x\} = \int_{-\infty}^x \frac{2}{\sigma} \cdot \varphi(t \cdot \sigma^{-1}) \cdot [1 - \Phi(t \cdot \lambda \cdot \sigma^{-1})] dt \qquad (9)$$

Where, $\sigma^2 = \sigma_o^2 + \sigma_u^2 (\dot{o} \sim \mathcal{N}(0,\sigma_o)$ and $u \sim \mathcal{N}^+(0,\sigma_u))$, $\lambda = \sigma_u / \sigma_o$, φ and $\ddot{O}$ stand for the density and the distribution functions of a standard normal law, respectively. Then, we have that

$$\mathcal{P}_0 = \mathcal{P}\{\bar{\varepsilon} \le 0\} = \mathcal{P}\{\dot{o} - u \le -\mathbb{E}[u]\}$$

$$= \int_{-\infty}^{-\sqrt{(2/\pi) \cdot \lambda}} \frac{2}{\sqrt{1+\lambda^2}} \cdot \varphi\left(\frac{t}{\sqrt{1+\lambda^2}}\right) \cdot \left[\ddot{O} - \left(\frac{\lambda \cdot t}{\sqrt{1+\lambda^2}}\right)\right] dt \in \mathcal{G} \lambda$$

i.e., $\mathcal{P}_0$ only depends on the λ values. In addition, $N \cdot \widehat{\mathcal{P}}_N^0$ follows again approximately a $\mathcal{B}(N, \mathcal{G}(\lambda))$ distribution, if $I_{N,\alpha/2}$ and $S_{N,\alpha/2}$ are the lower and the upper bound of a confidence interval (at level $1-\alpha$) for $\mathcal{G}(\lambda)$, then $\mathcal{G}^{-1}(I_{N,\alpha/2})$ and $\mathcal{G}^{-1}(S_{N,\alpha/2})$ are the upper and the lower bound of a confidence

interval (at level) for λ. Of course, bootstrap methods can also be used in order to improve the variance estimation.

Unfortunately, the function $\mathcal{G}^{-1}$ cannot be directly computed and, with this goal, numerical methods must be used. Figure 1 depicts the numeric computation for the function $\mathcal{G}$. Note that it is not defined for $\mathcal{P}\{\overline{\varepsilon} \leq 0\} \geq 1/2$.

Figure 1: Numeric computation for $G(\lambda), \lambda \in (0, 40)$.
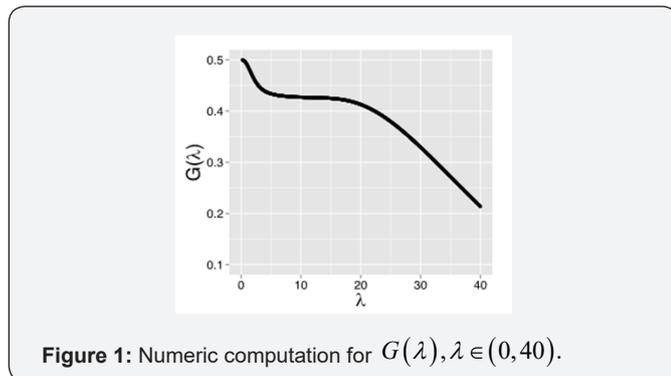
As it is well-known, the value of λ can also be derived from the $\mathcal{M}_3$ parameter. Particularly, it holds the following relationships:

$$\mathbb{V}[u] = \left(\frac{\mathcal{M}_3}{\sqrt{2/\pi}(1 - 4/\pi)} \right)^{2/3}$$

$$\mathbb{V}[\dot{o}] = \mathbb{V}[\varepsilon] - \frac{\pi - 2}{\pi} \cdot \mathbb{V}[u].$$

Therefore, for $\widehat{\mathcal{M}_3} \leq 0, \lambda$ can be directly estimated by using a plug-in method.

Monte carlo simulations

The practical behavior of the methodology exposed above is studied via Monte Carlo simulations. Three different scenarios have been considered. The first one is a simple normal-half normal situation; a sample with size N, X, is drawn from a standard normal distribution and then, it is computed $Y = 1 + X + \dot{o} - u$, where $\dot{o}$ follows a standard normal distribution and $u \sim \mathcal{N}^+(0, \lambda)$ with $\lambda^2 = 0$ (null), $0.5, 1, 2, 4$. The second scenario, contaminated normal-half normal, is similar to the previous one but, approximately, the 2.5%of the data are contaminated by a perturbation following the mixture $1/2 \cdot \mathcal{N}(-3, 0.1) + 1/2 \cdot \mathcal{N}(3, 0.1)$. Note that, in this scenario, the distribution of the white noise, $\dot{o}$ is not exactly a $\mathcal{N}(0,1)$ but, approximately, the mixture $0.0125 \mathcal{N}(-3, 0.1) + 0.975 \cdot \mathcal{N}(0,1) + 0.0125 \cdot \mathcal{N}(3, 0.1)$. Finally, in the third scenario, we consider a model with one dependent variable and three regressors the covariance matrix and the linear regression model structure as well as the total variance of the residuals are based on the WHO data adding a normal-half normal perturbation (complete information of this dataset is provided as Appendix).

Hypothesis testing

In this subsection, we investigate the capacity of the previous tests for detecting the presence of skewness under the three scenarios described above. Results based on 5,000 Monte Carlo iterations for the tests based on regions (5) (P_0) and (8) (S_k) , by using the standard Wald-based method, W(the R package frontier was used with this goal) and by a simple analysis of the residuals normality using the traditional Shapiro-Wilks test $(S - W)$ are provided in the following tables and figures. Table 1 shows the rejection proportions for the normal-half normal model when $\lambda = 0$ (null) at nominal levels $\alpha = 0.05$ and $\alpha = 0.10$. Nominal levels were respected by P_0, S_k and $(S - W)$ methods but, the standard W test obtained catastrophic results; at level $\alpha = 0.05$ between 12.8% and 21.5% of samples were rejected.

Table 1: Observed rejection proportions in 5,000Monte Carlo simulations for the normal-half normal model with $\lambda = 0$ (null hypothesis).

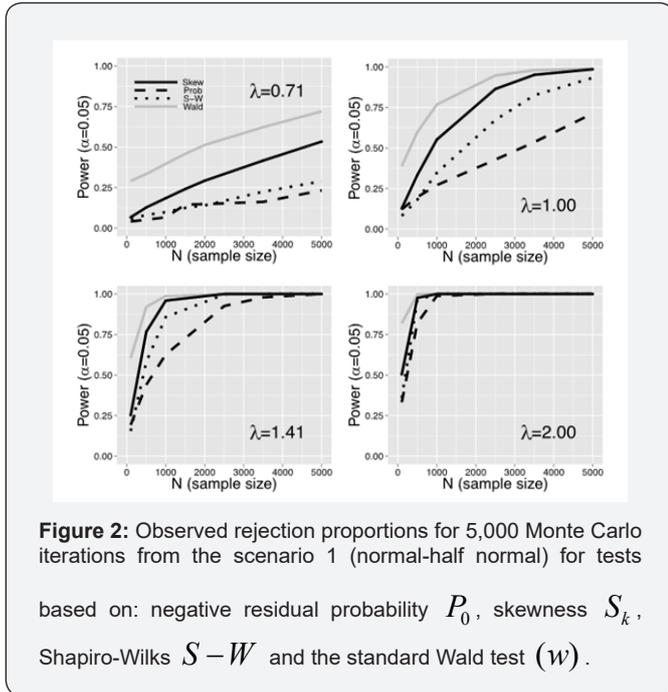| N	$\alpha = 0.05$				$\alpha = 0.10$			
	P_0	S_k	$S - W$	W	P_0	S_k	$S - W$	W
100	0.067	0.043	0.048	0.215	0.117	0.089	0.099	0.241
500	0.054	0.05	0.047	0.166	0.096	0.094	0.096	0.198
1000	0.05	0.045	0.05	0.151	0.094	0.096	0.099	0.196
2500	0.044	0.049	0.051	0.14	0.101	0.099	0.102	0.179
5000	0.052	0.05	0.041	0.128	0.1	0.083	0.086	0.17

Figure 2: Observed rejection proportions for 5,000 Monte Carlo iterations from the scenario 1 (normal-half normal) for tests based on: negative residual probability P_0, skewness S_k, Shapiro-Wilks $S-W$ and the standard Wald test (w).

Figure 2 shows the observed rejection proportions for the four considered tests for $\lambda^2 = 1/2, 1, 2$ and 4. Logically, the Wald test is always the best one; it obtained a good power even for $\lambda = 0$. It is followed by S_k and $S-W$. The new introduced test, P_0, obtained the worst results. Note that even for $\lambda = 1$ and N=1,000 the observed power does not achieve 80%. Table 2 shows the observed rejection proportions for 5,000 Monte Carlo iterations under the scenario 2 (contaminated normal-half normal). All considered test overestimated the nominal level. Logically, Shapiro-Wilks test increases the rejection proportion with the sample size; note that, in this case, the null hypothesis is not true because the residuals are normal distributed. However, test based on P_0 shows itself more robust than the others; at level $\alpha = 0.05$, it rejected a maximum of 8.3% (N=100) and around the 6% when sample size increases. The rejection proportion for the tests based on S_k and W increased with the sample size to achieve the 20.3% and 60.5% for N=5,000 and $\alpha = 0.05$, respectively.

Table 2: Observed rejection proportions in 5,000 Monte Carlo simulations for the contaminated normal-half normal model with $\lambda = 0$ (null hypothesis).

$\alpha = 0.05$					$\alpha = 0.10$			
N	P_0	S_k	$S-W$	W	P_0	S_k	$S-W$	W
100	0.083	0.167	0.344	0.308	0.142	0.227	0.423	0.337
500	0.066	0.188	0.846	0.354	0.116	0.244	0.89	0.344
1000	0.066	0.208	0.981	0.407	0.117	0.258	0.99	0.443
2500	0.062	0.211	1	0.521	0.116	0.256	1	0.559
5000	0.063	0.203	1	0.605	0.109	0.255	1	0.643

Table 3: Observed rejection proportions in 5,000 Monte Carlo simulations for the WHO based data with normal-half normal residuals for $\lambda = 0$ (null hypothesis).

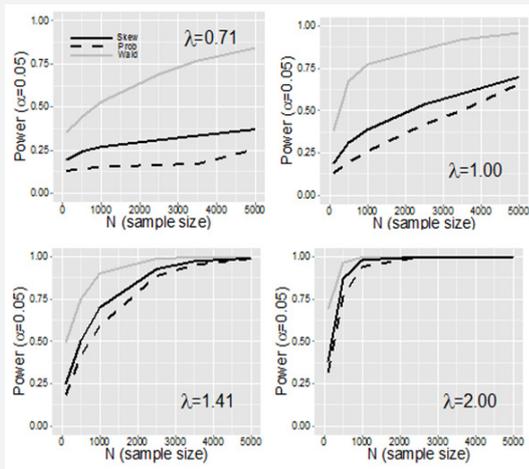| $\alpha = 0.05$					$\alpha = 0.10$			
N	P_0	N	P_0	N	P_0	N	P_0	N
100	0.083	0.167	0.344	0.308	0.142	0.227	0.423	0.337
500	0.066	0.188	0.846	0.354	0.116	0.244	0.89	0.344
1000	0.066	0.208	0.981	0.407	0.117	0.258	0.99	0.443
2500	0.062	0.211	1	0.521	0.116	0.256	1	0.559
5000	0.063	0.203	1	0.605	0.109	0.255	1	0.643

Figure 3: Observed rejection proportions for 5,000 Monte Carlo iterations from the scenario 2 (contaminated normal-half normal) for tests based on: negative residual probability P_0, skewness S_k, Shapiro-Wilks $S-W$ and the standard Wald test (w).
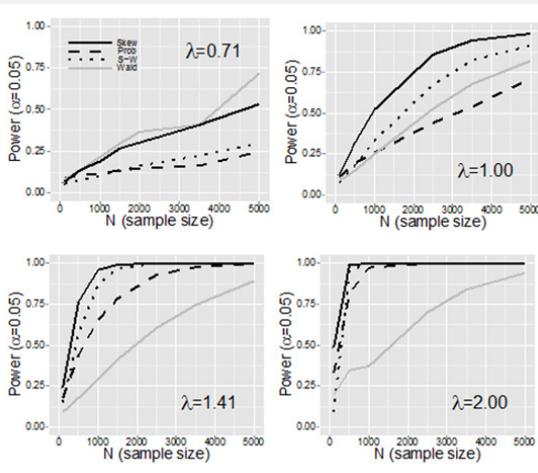

Figure 4: Observed rejection proportions for 5,000 Monte Carlo iterations from the scenario 3 (structure of the data extracted from the WHO dataset) for tests based on: negative residual probability P_0, skewness S_k, Shapiro-Wilks $S-W$ and the standard Wald test (w).

Figure 3 is similar to Figure 2 for the contaminated normal-half normal scenario. Because, at this scenario, it does not hold the residuals normality even for $\ddot{\mathbf{e}} = \mathbf{0}$, the Shapiro-Wilks test has not been included in this plot. The Wald test was, again, the most powerful one; remember that the rejection proportion under the null was also really high. Tests based on S_k and P_0 had similar behaviour although S_k was always better than P_0. Finally, Table

3 shows the observed rejection proportions, under the null, for the third considered scenario. Tests based on P_0, S_k and $S-W$ respected the two considered nominal levels. However, the Wald test increased the rejection proportion with the sample size, even it overcame the 50% for N=5,000and α=0.05. Figure 4 shows the observed power for the four studied tests in the third described scenario. The results were not essentially different to the ones observed in Figure 2. However, it is worth to remark that, in this scenario which is more complex than the previous ones, for $\lambda^2 = 1/2, 1, 2$, the Wald test worked worse than S_k and, for the largest λ, even worse than the Shapiro-Wilks and P_0 tests. In addition, P_0 shows itself more competitive since it is not affected for the new scenario.

Confidence intervals

In order to study the quality of the estimation provided by the skewness based methods and the usual maximum-likelihood based one, we explored the coverage percentage of symmetric 95% confidence intervals built by using the usual naive bootstrap method (200 replications) for the procedures based on the probability (P_0) and (S_k) the moments . In addition, the R package frontier is used to compute the confidence interval for the maximum-likelihood method. Table 4 shows the observed results in 5,000 Monte Carlo iterations from the scenario 1. Obviously, estimations have only been computed when it was possible (final sample used is provided as % valid). With the coverage percentage, the length of the 95% confidence interval is reported by $mean \pm sd \, (standard \; deviation)$. Results show that only the P_0 estimator achieved the expected coverage percentage in all situations. S_k an W_d estimators obtained really poor results for small sample and mild $\lambda-values$. The W estimator directly cannot be used for λ^2 lower than 1 and if $\lambda^2 = 1$ sample size must be really large. Table 5 is similar to Table 4 when samples are drawn from scenario 2. The three considered tests obtained bad results in this setting. S_k and Westimators achieved really poor coverage. These results were only similar to the expected ones for λ=2. In addition, the coverage percentage decreased with the sample size for the three investigated estimators. In spite of the length of the respective confidence intervals were the largest, only P_0 obtained reasonable results for most of the considered situations. Table 6 shows the observed results for the scenario 3. These results were not different to the ones obtained in Table 4. Usual estimator based on the maximum-likelihood ratio only worked for $\lambda^2 \geq 2$ and $N \geq 1,000$.

Table 4: Coverage percentage for the 95% confidence interval and length of these confidence intervals expressed as $mean \pm sd$ (standard deviation) for the estimators based on P_0, S_k (confidence intervals computed via 200bootstrap resamples) and W estimators (computed with the R package frontier) for the scenario 1 (normal-half normal).

			% Coverage			Conf. Int. length($mean \pm sd$)		
λ^2	N	% valid	P_0	S_k	W	P_0	S_k	W
1/2	500	8.50%	0.939	0.628	0.143	2.72 ± 0.78	1.42 ± 0.35	1.82 ± 0.31
	1000	13.00%	0.931	0.777	0.401	1.85 ± 0.22	1.04 ± 0.18	1.14 ± 0.12
	5000	40.00%	0.96	0.944	0.876	1.10 ± 0.16	0.59 ± 0.08	0.56 ± 0.06
1	500	24.80%	0.958	0.911	0.675	2.83 ± 1.17	1.47 ± 0.47	1.80 ± 0.30
	1000	40.80%	0.974	0.942	0.852	1.87 ± 0.23	1.03 ± 0.18	1.10 ± 0.11
	5000	95.50%	0.967	0.973	0.957	1.01 ± 0.22	0.49 ± 0.007	0.48 ± 0.04
2	500	66.50%	0.985	0.978	0.897	3.30 ± 2.14	1.57 ± 0.76	1.75 ± 0.27
	1000	91.90%	0.979	0.976	0.935	1.02 ± 0.29	1.02 ± 0.22	1.06 ± 0.09
	5000	99.70%	0.954	0.944	0.959	0.72 ± 0.09	0.44 ± 0.05	0.42 ± 0.03
4	500	97.60%	0.984	0.949	0.931	5.61 ± 4.71	2.10 ± 1.70	1.88 ± 0.35
	1000	100%	0.962	0.93	0.943	2.33 ± 1.53	1.30 ± 0.58	1.12 ± 0.23
	5000	100%	0.957	0.94	0.95	0.77 ± 0.08	0.54 ± 0.08	0.45 ± 0.09

Table 5: Coverage percentage for the 95% confidence interval and length of these confidence intervals expressed as $mean \pm sd$ (standard deviation) for the estimators based on P_0, S_k (confidence intervals computed via 200bootstrap resamples) and W estimators (computed with the R package frontier) for the scenario 2 (contaminated normal-half normal).

			% Coverage			Conf. Int. length($mean \pm sd$)		
λ^2	N	% valid	P_0	S_k	W	P_0	S_k	W
1/2	500	43.20%	0.922	0.338	0.02	2.90 ± 1.31	3.04 ± 8.27	1.41 ± 0.23
	1000	70.00%	0.898	0.283	0.034	1.90 ± 0.25	1.54 ± 068	0.94 ± 0.09
	5000	99.20%	0.646	0.005	0.001	0.95 ± 0.22	0.60 ± 0.08	0.39 ± 0.02
1	500	58.80%	0.926	0.697	0.358	3.18 ± 2.01	2.51 ± 2.05	1.58 ± 0.23
	1000	86.20%	0.924	0.592	0.392	1.93 ± 0.24	1.37 ± 0.58	0.97 ± 0.10
	5000	99.80%	0.733	0.046	0.038	0.8 ± 016	0.56 ± 0.08	0.39 ± 0.04

	500	85.30%	0.965	0.915	0.752	3.85 ± 2.99	2.23 ± 1.78	1.67 ± 0.25
2	1000	98.50%	0.948	0.829	0.718	1.98 ± 0.48	1.32 ± 069	1.01 ± 0.10
	5000	100%	0.814	0.293	0.328	0.71 ± 0.06	0.55 ± 0.09	0.41 ± 0.06
	500	98.80%	0.976	0.965	0.876	6.58 ± 5.38	2.77 ± 2.52	1.88 ± 0.37
4	1000	99.90%	0.956	0.931	0.868	2.69 ± 2.11	1.66 ± 1.28	1.14 ± 014
	5000	100%	0.912	0.799	0.768	0.82 ± 0.10	0.64 ± 0.13	0.46 ± 0.03

Table 6: Coverage percentage for the 95% confidence interval and length of these confidence intervals expressed as $mean \pm sd$ (standard deviation) for the estimators based on P_0, S_k (confidence intervals computed via 200bootstrap resamples) and W estimators (computed with the R package frontier) for the scenario 3 (normal-half normal based on the WHO data structure).

			% Coverage			Conf. Int. length($mean \pm sd$)		
λ^2	N	% **valid**	P_0	S_k	W	P_0	S_k	W
	500	8.50%	0.946	0.686	0.163	2.67	1.43	1.88 ± 0.35
1/2	1000	12.40%	0.934	0.779	0.389	1.88	1.03	1.15 ± 0.11
	5000	39.90%	0.958	0.944	0.881	1.12	0.59	0.56 ± 0.06
	500	24.20%	0.977	0.91	0.636	2.83	1.48	1.81 ± 0.32
1	1000	41.10%	0.972	0.951	0.837	1.9	1.04	1.11 ± 0.11
	5000	95.90%	0.972	0.972	0.949	1.01	0.49	0.48 ± 0.04
	500	64.60%	0.988	0.978	0.882	3.33	1.59	1.78 ± 0.59
2	1000	91.30%	0.988	0.975	0.932	1.45	1.02	1.06 ± 0.09
	5000	99.90%	0.961	0.942	0.947	0.72	0.44	0.41 ± 0.01
	500	99.80%	0.97	0.925	0.864	2.36	1.28	1.13 ± 0.11
4	1000	99.90%	0.956	0.931	0.94	1.69	0.84	0.73 ± 0.04
	5000	100%	0.954	0.941	0.942	0.78	0.54	0.45 ± 0.02

Real data analysis: the WHO data

In order to illustrate the practical behavior of the investigated estimators, the well-known data from the World Health Organization (WHO) is studied [17]. Particularly, we consider data referred to the year 1997 with a total of 227 world regions $(= N$ and we try to model the Disability Adjusted Life Expectancy (DALE) from the regressors: per pro-capita health expenditure (HEXP) and educational attainment (HC). Last variable is considered in quadratic form and therefore the regressor HC2 (square of the educational attainment) is also included in the model. Since the logarithm transformation is usually applied, the standard least square error method provides the equation,

$$\log\log(DALE) = 3.24 + 0.07 \cdot \log(HEXP) + 0.25 \cdot \log(HC) - 0.03 \cdot \log(HC2) + \varepsilon$$

Where, $\mathbb{V}[\hat{\sigma}] = 0.017$ and $R^2 = 0.719$. $\mathcal{P}_0 = \mathcal{P}\{\varepsilon \leq 0\} = 0.493$, hence based on the probability estimator, and assuming the normal-half normal model $\hat{\lambda} = 0.874$; the *p-value* associated with the null $H_0 : \mathcal{P}_0 = 1/2$ was 0.629. On the other hand, $\widehat{\mathcal{M}_3} = -0.003$; the *p-value* associated with the null $H_0 : \mathcal{M}_3 = 0$ was 0.929. The value of $\widehat{\mathcal{M}_3}$ is too close to zero and does not allow to estimate $\lambda > 0$. The maximum-likelihood based procedure provided a λ

estimator of 5.76 with a p-value=0.001. The parameters were also slightly different and the obtained equation was,

$$\log(DALE) = 3.47 + 0.06 \cdot \log(HEXP) + 0.37 \cdot \log(HC) - 0.15 \cdot \log(HC2) + \dot{o} - u$$

Figure 5 depicts the densities for the model based on the maximum-likelihood with residuals normal-half normal distributed $\left(\hat{\lambda} = 5.76\right)$, on $\widehat{\mathcal{P}}_N^0\left(\hat{\lambda} = 0.874\right)$, for normal residuals $(\lambda = 0)$ and kernel density estimation for the observed residuals. Although it seems that the residuals are not drawn from a normal law Shapiro-Wilks test rejects this hypothesis with a p-value=0.001, the kernel density estimation (KDE) is far from the normal model but also far from the normal-half normal estimated from the maximum-likelihood method.
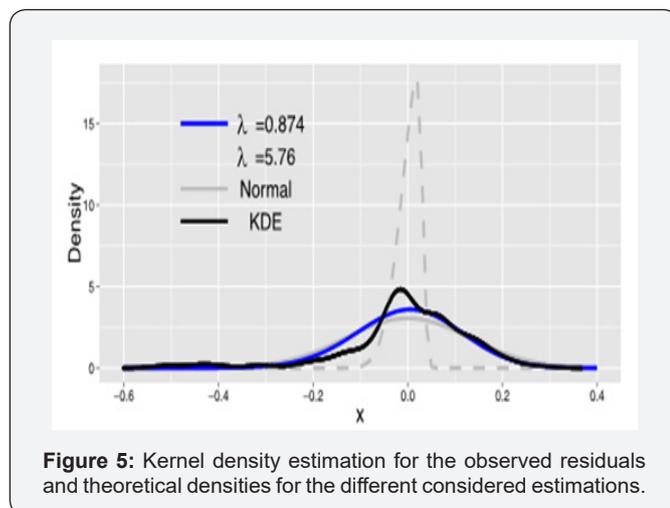
Figure 5: Kernel density estimation for the observed residuals and theoretical densities for the different considered estimations.

Discussion

Stochastic frontier models are conventionally used in order to study production and cost functions. The specialized literature focuses are, mainly, the estimation of the parameters of these functions by regression techniques. Parametric normal-half normal residuals joint with the maximum-likelihood method is, probably, the most frequently employed procedure. In the voluminous literature on this topic, statistical testing of the properties of the disturbance term has attracted little attention. In this work, the authors deal with the stochastic frontier model identification from a statistical approach. Via Monte Carlo simulations, we study the performance of the Wald test computed from the R package frontier. The results were catastrophic; Wald test respects neither the nominal levels nor the coverage percentages in most of the considered situations. Moreover, in the outliers' presence, the obtained results were extremely poor. These results support previous works which achieved similar conclusions [11,14].

Similarly, to Kuosmanen & Fosgerau [15], we proposed to study the residuals expected skewness in order to make inference about the data capacity to establish a stochastic frontier model. In spite that Simar and Wilson [16] and Feng et al. [18] worked on the so-called "wrong" skewness and it seems clear that, even when the stochastic frontier model is correctly specified, data can present "wrong" skewness. We assume that, in these cases, data have not provided enough evidence to estimate the potential SFA.

The new proposed estimator, based on the expression $\mathcal{P}\left\{\overline{\varepsilon} \leq 0\right\}$ is really easy to compute, fully non-parametric and very robust, which implies being a bit conservative when there is not outlier in the sample. The analysis of the considered real dataset was really enlightening; while $\widehat{\mathcal{M}_3}$ estimator provided a really close to zero estimation, $\widehat{\mathcal{P}}_N^0$ obtained a λ estimation close to 1 (it was poor taking into account the available sample size and the power observed in the Monte Carlo study). The Wald method provides a $\lambda - value$ larger than 5. The observation of the residuals distribution suggests a lack of normality in their distribution, but the adjustment to the normal-half normal model proposed by the maximum-likelihood method was even worse. It seems that the maximum-likelihood method assigns all deviation from normality to the so-called inefficiency; however, as it is well-known, a regression model can have different problems which produce a lack of normality in the residuals distribution.

In short, stochastic frontier models frequently try to detect a slight deviation in the normality of the residuals; this deviation is assigned to the inefficiency term. However, detecting these deviations is really difficult without huge samples. In addition, the origin of these deviations cannot be the inefficiency term; in particular, there exists different causes which provoke this effect. A deep study of the residuals distribution, particularly of their symmetry can drive to obtain a good knowledge about the real nature of the model. Of course, the application of (relatively) modern techniques to analyzed the symmetry of the perturbance distributions or methods based on quantiles among many others can be useful [19,20].

Acknowledgement

References

1. Aigner D, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontiers production function models. Journal of Econometrics 6: 21-37.

2. Meeusen W, Van den BJ (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. International Economic Review 18: 435-444.

3. Fan Y, Li Q, Weersink A (1996) Semiparametric estimation of stochastic production frontier models. Journal of Business and Economic Statistics 14: 460-468.

4. Martins-FC, Yaoc F (2015) Semiparametric Stochastic Frontier Estimation Via Profile Likelihood. Econometric Reviews 34(4): 413-451.

5. Kumbhakar SC, Park BU, Simar L, Tsionas LG (2007) Nonparametric stochastic frontiers: A local likelihood approach. Journal of Econometrics 137(1): 1-27.

6. Greene WH (1993) The econometric approach to efficiency analysis, In: Fried HO, Lovell CAK, Schmidt SS (Eds), The Measurement of Productive Efficiency, Oxford University Press, New York, USA, pp. 68-119.

7. Stevenson RE (1980) Likelihood function for generalized stochastic frontier estimation. Journal of Econometrics, 13: 57-66.

8. Greene WH (1990) A gamma-distributed stochastic frontier model. Journal of Econometrics 46: 141-163.

9. Schmidt P, Lovell AK (1979) Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. Journal of Econometrics 13: 83-100.

10. Simar L, Van Keilegom I, Zelenyuk V (2014) Nonparametric least squares methods for stochastic frontier models, Working paper series; Centre for Efficiency and Productivity Analysis.

11. Coelli TJ (1995) Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis. Journal of Productivity Analysis 6: 247-268.

12. Olson JA, Schmidt P, Waldman DM (1980) A Monte Carlo study of estimators of stochastic frontier production functions. Journal of Econometrics 13: 67-82.

13. Behr A, Tente S (2008) Stochastic frontier analysis by means of maximum likelihood and the method of moments, Discussion Paper Series 2: Banking and Financial Studies.

14. Simar L, Wilson PW (2009) Inferences from cross-sectional stochastic frontier models. Econometric Reviews 29(1): 62-98.

15. Kuosmanen T, Fosgerau M (2009) Neoclassical versus frontier production models? Testing for the skewness of Regression Residuals. Scandinavian Journal of Economics 111(2): 351-367.

16. Wilson PW (2008) FEAR: A software package for frontier efficiency analysis with R, Socio-Economic Planning Sciences 42(4): 247-254.

17. Evans D, Tandon A, Murray C, Lauer J (2000) The comparative efficiency of national health systems in producing health: an analysis of 191 countries. World Health Organization, GPE Discussion pp. 29.

18. Feng Q, Horrace WC, Wu GL (2013) Wrong skewness and finite sample correction in parametric stochastic frontier models, Center for Policy Research Maxwell School of Citizenship and Public Affairs Syracuse University, USA.

19. Ahmad IA, Li Q (1997) Testing symmetry of an unknown density function by kernel method, Nonparametric Statistics 7(3): 279-293.

20. Staudte RG (2014) Inference for quantile measures of skewness. Test 23(4): 751-768.